Descriptors for uploading information from National Inventories to EURISCO

Introduction

This descriptor list is used for uploading information from the National Inventories to EURISCO, and thus purely a format of data exchange.

The list is an extension of the FAO/IPGRI multi-crop passport descriptors (MCPD) which were published December 2001, developed jointly by IPGRI and FAO, with input from many documentation specialists worldwide, to provide international standards to facilitate germplasm passport information exchange. All MCPD are included, without change and with the same format rules, in the current list. Eight descriptors (0, 29-35) were added for the specific purposes of EURISCO: the first descriptor, identifying the National Inventory, five allowing the incorporation of information relevant to EURISCO, which otherwise would not fit in the MCPD and since 2008 other two were added (MLS, AEGIS).

General format rules

Following format rules, as copied from the MCPD-list, apply to all fields:

- If a field allows multiple values, these values should be separated by a semicolon (;) without space(s). (i.e. Accession name: "Rheinische Vorgebirgstrauben;Emma;Avlon")
- A field for which no value is available should be left empty (i.e. Elevation). If data are exchanged in ASCII format for a field with a missing numeric value, it should be left empty. If data are exchanged in a database format, missing numeric values should be represented by generic NULL values.
- Dates are recorded as YYYYMMDD. If the month and/or day are missing this should be indicated with hyphens. Leading zeros are required (i.e. 197506--, or 1975----).
- Latitude and longitude are recorded in an alphanumeric format. If the minutes or seconds are missing, this should be indicated with hyphens. Leading zeros are required.
- For coding countries three-letter ISO 3166-1 codes are used (including the codes that are no longer in use in the ISO 3166-1, such as DDR).[1]
- For coding institutes the FAO Institute Codes should be used as maintained by the FAO. The codes consist of the 3-letter ISO 3166 country code of the country where the institute is located plus a three-digit number.[2]
- The preferred language for free text fields is English (i.e. Location of collecting site and Remarks).

Descriptors

The descriptors are numbered according to the FAO/IPGRI multi-crop passport descriptors (MCPD); the first descriptor (numbered 0) and the last seven (numbered 29-35) are additional, and specific to this EURISCO Descriptor List. Only the four fields identifying the accession are mandatory, all other fields are highly recommended. The mandatory fields are NICODE (0), INSTCODE (1), ACCENUMB (2) and GENUS (5). The combination of these fields has to be unique.

---

[1] The ISO 3166-1 Code List can be found at: http://unstats.un.org/unsd/methods/m49/m49alpha.htm. Country or area numerical codes added or changed are not available on line at: http://unstats.un.org/unsd/methods/m49/m49chang.htm.

[2] These codes are available from http://apps3.fao.org/wiews/wiews.jsp for registered WIEWS users. From the Main Menu select: 'PGR' and 'Download'. If new Institute Codes are required, they can be generated online by national WIEWS correspondents, or by the FAO WIEWS administrator [Stefano.Diulgheroff@fao.org].

MCPD + EURISCO DESCRIPTORS

0. National Inventory code (NICODE)

Code identifying the National Inventory; the code of the country preparing the National Inventory. Exceptions are possible, if agreed with EURISCO such as NGB.

Example: `NLD`

1. Institute code (INSTCODE)

FAO Institute Code of the institute where the accession is maintained.

Example: `NLD037`

2. Accession number (ACCENUMB)

This number serves as a unique identifier for accessions within a genebank collection, and is assigned when a sample is entered into the genebank collection.

Example: `CGN00254`

3. Collecting number (COLLNUMB)

Original number assigned by the collector(s) of the sample, normally composed of the name or initials of the collector(s) followed by a number. This number is essential for identifying duplicates held in different collections.

Example: `FA90-110`

4. Collecting institute code (COLLCODE)

Code of the Institute collecting the sample. If the holding institute has collected the material, the collecting institute code (COLLCODE) should be the same as the holding institute code (INSTCODE).

Example: `NLD037`

5. Genus (GENUS)

Genus name for taxon, in latin. Initial uppercase letter required.

Example: `Allium`

6. Species (SPECIES)

Specific epithet portion of the scientific name, in latin, in lowercase letters. Following abbreviation is allowed: 'sp.'

Example: `paniculatum`

7. Species authority (SPAUTHOR)

The authority for the species name.

Example: `L.`

8. Subtaxa (SUBTAXA)

Subtaxa can be used to store any additional taxonomic identifier, in latin. Following abbreviations are allowed: 'subsp.' (for subspecies); 'convar.' (for convariety); 'var.' (for variety); 'f.' (for form).

Example: `subsp. fuscum`

9. Subtaxa authority (SUBTAUTHOR)

The subtaxa authority at the most detailed taxonomic level.

Example: `(Waldst. et Kit.) Arc.`

10. Common crop name (CROPNAME)

Name of the crop in colloquial language, preferably English.

Example: `malting barley`
Example: `cauliflower`

11. Accession name (ACCENAME)

Either a registered or other formal designation given to the accession. First letter uppercase. Multiple names separated with semicolon without space.

Example: `Rheinische Vorgebirgstrauben;Emma;Avlon`

12. Acquisition date (ACQDATE)

Date on which the accession entered the collection as YYYYMMDD. Missing data (MM or DD) should be indicated with hyphens. Leading zeros are required.

Example: `1968----`
Example: `20020620`

13. Country of origin (ORIGCTY)

Code of the country in which the sample was originally collected.

Example: `NLD`

14. Location of collecting site (COLLSITE)

Location information below the country level that describes where the accession was collected. This might include the distance in kilometres and direction from the nearest town, village or map grid reference point

Example: `7 km south of Curitiba in the state of Parana`

15. Latitude of collecting site* (LATITUDE)

Degree (2 digits) minutes (2 digits), and seconds (2 digits) followed by N (North) or S (South). Every missing digit (minutes or seconds) should be indicated with a hyphen. Leading zeros are required

Example: `10----S`
Example: `011530N`
Example: `4531--S`

16. Longitude of collecting site* (LONGITUDE)

Degree (3 digits), minutes (2 digits), and seconds (2 digits) followed by E (East) or W (West). Every missing digit (minutes or seconds) should be indicated with a hyphen. Leading zeros are required.

Example: `0762510W`
Example: `076----W`

17. Elevation of collecting site (ELEVATION)

Elevation of collecting site expressed in meters above sea level. Negative values are allowed.

Example: `763`

18. Collecting date of sample (COLLDATE)

Collecting date of the sample as YYYYMMDD. Missing data (MM or DD) should be indicated with hyphens. Leading zeros are required.

Example: `1968----`
Example: `20020620`

19. Breeding institute code (BREDCODE)

FAO Institute Code of the institute that has bred the material.

20. Biological status of accession (SAMPSTAT)

The coding scheme proposed can be used at 3 different levels of detail: either by using the general codes (in boldface) such as 100, 200, 300, 400 or by using the more specific codes such as 110, 120 etc.

- 100) Wild
    - 110) Natural
    - 120) Semi-natural/wild
- 200) Weedy
- 300) Traditional cultivar/landrace
- 400) Breeding/research material
    - 410) Breeder's line
        - 411) Synthetic population
        - 412) Hybrid
        - 413) Founder stock/base population
        - 414) Inbred line (parent of hybrid cultivar)
        - 415) Segregating population
    - 420) Mutant/genetic stock
- 500) Advanced/improved cultivar
- 999) Other (Elaborate in REMARKS field)

* To convert from longitude and latitude in degrees (º), minutes ('), seconds ("), and a hemisphere (North or South and East or West) to decimal degrees, the following formula should be used:

$$d° \, m' \, s'' = h *(d + m/60 + s/3600)$$

where h=1 for the Northern and Eastern hemispheres and –1 for the Southern and Western hemispheres i.e. 30º30'0'' S= –30.5 and 30º15'55'' N=30.265.

21. Ancestral data	**(ANCEST)**

Information about either pedigree or other description of ancestral information (i.e. parent variety in case of mutant or selection).

Example: `Hanna/7*Atlas//Turk/8*Atlas`
Example: `mutation found in Hanna`
Example: `selection from Irene`
Example: `cross involving amongst others Hanna and Irene`

22. Collecting/acquisition source	**(COLLSRC)**

The coding scheme proposed can be used at 2 different levels of detail: either by using the general codes (in boldface) such as 10, 20, 30, 40 or by using the more specific codes such as 11, 12 etc.

- **10) Wild habitat**
  - 11) Forest/woodland
  - 12) Shrubland
  - 13) Grassland
  - 14) Desert/tundra
  - 15) Aquatic habitat
- **20) Farm or cultivated habitat**
  - 21) Field
  - 22) Orchard
  - 23) Backyard, kitchen or home garden (urban, peri-urban or rural)
  - 24) Fallow land
  - 25) Pasture
  - 26) Farm store
  - 27) Threshing floor
  - 28) Park
- **30) Market or shop**
- **40) Institute, Experimental station, Research organization, Genebank**
- **50) Seed company**
- **60) Weedy, disturbed or ruderal habitat**
  - 61) Roadside
  - 62) Field margin
- **99) Other (Elaborate in REMARKS field)**

23. Donor institute code	**(DONORCODE)**

FAO Institute Code for the donor institute.

24. Donor accession number	**(DONORNUMB)**

Number assigned to an accession by the donor.

Example: `NGB1912`

25. Other identification (numbers) associated with the accession	**(OTHERNUMB)**

Any other identification (numbers) known to exist in other collections for this accession. Use the following system: INSTCODE:ACCENUMB;INSTCODE:ACCENUMB;… INSTCODE and ACCENUMB follow the standard described above and are separated by a colon. Pairs of INSTCODE and ACCENUMB are separated by a semicolon without space. When the institute is not known, the number should be preceded by a colon.

Example: `NLD037:CGN00254`
Example: `SWE002:NGB1912;:Bra2343`

26. Location of safety duplicates	**(DUPLSITE)**

FAO Institute Code of the institute where a safety duplicate of the accession is maintained. The codes consist of the 3-letter ISO 3166 country code of the country where the institute is located plus a number.

27. Type of germplasm storage	**(STORAGE)**

If germplasm is maintained under different types of storage, multiple choices are allowed (separated by a semicolon). (Refer to FAO/IPGRI Genebank Standards 1994 for details on storage type.)

- **10) Seed collection**
  - 11) Short term
  - 12) Medium term
  - 13) Long term
- **20) Field collection**
- **30) In vitro collection (Slow growth)**
- **40) Cryopreserved collection**
- **99) Other (elaborate in REMARKS field)**

28. Remarks	**(REMARKS)**

The remarks field is used to add notes or to elaborate on descriptors with value 99 or 999 (=Other). Prefix remarks with the field name they refer to and a colon. Separate remarks referring to different fields are separated by semicolons without space.

Example: `COLLSRC:roadside`

29. Decoded collecting institute (COLLDESCR)

Brief name and location of the collecting institute. Only to be used if COLLCODE can not be used since the FAO Institution Code for this institute is not (yet) available.

Example: `Tuinartikelen Jan van Zomeren, Arnhem, The Netherlands`

30. Decoded breeding institute (BREDDESCR)

Brief name and location of the breeding institute. Only to be used if BREDCODE can not be used since the FAO Institution Code for this institute is not (yet) available.

Example: `CFFR from Chile`

31. Decoded donor institute (DONORDESCR)

Brief name and location of the donor institute. Only to be used if DONORCODE can not be used since the FAO Institution Code for this institute is not (yet) available.

Example: `Nelly Goudwaard, Groningen, The Netherlands`

32. Decoded safety duplication location (DUPLDESCR)

Brief name and location of the institute maintaining the safety duplicate. Only to be used if DUPLSITE can not be used since the FAO Institution Code for this institute is not (yet) available.

Example: `Pakhoed Freezers inc., Paramaribo, Surinam`

33. Accession URL (ACCEURL)

URL linking to additional data about the accession either in the holding genebank or from another source.

Example: `www.cgn.wageningen-ur.nl/pgr/collections/passdeta.asp?accenumb=CGN04848`

34. MLS Status (MLSSTAT)

The coded status of an accession with regards to the Multilateral System (MLS) of the International Treaty on Plant Genetic Resources for Food and Agriculture.

Provides the information, whether the accession is included in the MLS.

0- not part of the MLS

1- part of the MLS.

If the MLS status is unknown, the field stays empty

35. AEGIS Status (AEGISSTAT)

The coded status of an accession with regards to the European Genebank Integrated System (AEGIS).

Provides the information, whether the accession is conserved for AEGIS.

0- not part of the AEGIS

1- part of the AEGIS.

If the AEGIS status is unknown, the field stays empty